Journal of Advances in Developmental Research (IJAIDR)

E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

Human Stress Analysis Using Machine Learning

K. Divya[1], S. Mari subbulakshmi[2]

[2]Assistant Professor
[1,2]Department of Computer science, Jeppiaar University, Chennai, India.

Abstract:
Stress has become a critical factor affecting human health, productivity, and overall well-being in modern society. Early detection and management of stress are essential to prevent severe mental and physical health issues such as anxiety, depression, and cardiovascular diseases. This paper presents a machine learning-based approach for analysing and detecting human stress levels using physiological and behavioural data. Features such as heart rate, skin temperature, electrodermal activity, facial expressions, and voice patterns are utilized to identify stress conditions. The collected data is pre-processed and analysed using supervised learning algorithms such as Support Vector Machine (SVM), Random Forest, and Decision Tree classifiers. The system categorizes stress levels into low, medium, and high. The proposed model aims to provide a non-invasive, real-time, and accurate stress detection mechanism. This approach can be applied in healthcare systems, workplace monitoring, and personal wellness applications, enabling timely intervention and improved mental health management.

Index Term: Machine Learning, Stress Detection, Human Stress Analysis, Physiological Signals, Support Vector Machine (SVM), Random Forest, Decision Tree, Feature Extraction, Data Preprocessing, Classification, Wearable Sensors, Mental Health Monitoring, Real-Time Prediction, Artificial Intelligence, Behavioural Analysis.

I. INTRODUCTION

Stress is a natural psychological and physiological response to challenging situations or demands. While moderate stress can enhance performance and motivation, prolonged or excessive stress can negatively impact both mental and physical health. It is associated with serious health conditions such as hypertension, anxiety disorders, depression, and heart diseases. Traditional stress detection methods primarily rely on self-assessment questionnaires, interviews, and clinical evaluations. These methods are often subjective, time-consuming, and unsuitable for continuous monitoring. With advancements in technology, machine learning has emerged as an effective tool for analysing large-scale data and identifying hidden patterns. Machine learning techniques enable automated stress detection by analysing physiological signals such as heart rate, skin conductivity, and behavioural indicators like facial expressions and speech patterns. This project aims to develop a robust and scalable system that classifies stress levels into categories such as low, medium, and high using supervised learning algorithms. The system can be integrated into wearable devices and mobile applications to provide real-time stress monitoring and management.

A. Abbrevations and Acronyms

AI (Artificial Intelligence), ML (Machine Learning), SVM (Support Vector Machine), RF (Random Forest), DT (Decision Tree), ANN (Artificial Neural Network), CNN (Convolutional Neural Network), RNN (Recurrent Neural Network), ECG (Electrocardiogram), EEG (Electroencephalogram), GSR (Galvanic Skin Response), HR (Heart Rate), BP (Blood Pressure), IoT (Internet of Things), EDA (Exploratory Data Analysis), and PCA (Principal Component Analysis).

Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

B. Units
Standard SI units are used for all measurements:
Beats per minute (bpm) is a measure of heart rate.
Temperature of the body in degrees Celsius (°C)
Hours of Sleep (hours)
Daily Steps of Physical Activity
Time: Minutes (min), Seconds (s)
Reliability and consistency are guaranteed when standard units are used.

C. Equations
The mathematical basis of machine learning models used for stress analysis is represented by equations. This paper's equations are all sequentially numbered and appropriately aligned using LaTeX settings. For instance, the classification model's accuracy is determined as follows: The formula for accuracy is \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions}. Similarly, model performance is assessed using precision and recall: Precision is equal to \frac{TP}{TP + FP}, and recall is equal to \frac{TP}{TP + FN}, where TP stands for true positives, FP for false positives, and FN for false negatives. These assessment measures aid in determining how well machine learning algorithms like Random Forest, Decision Tree, and Support Vector Machine (SVM) categorize stress levels. Every symbol used in an equation has a clear definition, either before or after the equation.

II. LITERATURE SURVEY
Several exploration workshops have been conducted in the field of stress discovery using machine literacy ways.
1. Zhang et al.(2019) proposed a deep literacy-based model that uses smartphone and wearable data to ameliorate stress vaccination delicacy by incorporating behavioural environment.
2. Hernandez et al.(2020) developed a multimodal stress discovery system combining physiological signals similar as heart rate variability(HRV), Possible spelling mistake found. exertion(EDA), and tone- reported data, enhancing system robustness.
3. Islam et al.(2021) introduced a mongrel deep literacy model using EEG and ECG signals, demonstrating better bracket performance compared to traditional machine literacy styles.
4. Sultana et al.(2022) employed wearable detector data with XG Boost algorithms to handle noisy physiological data effectively.
5. Baniya and Preminger (2023) stressed respiration rate as a significant index of stress using ECG and breathing signals.
6. Chen et al.(2024) proposed a real-time stress monitoring system using multimodal deep literacy ways, achieving high precision with temporal data analysis.

These studies demonstrate that combining multiple data sources and advanced machine literacy ways significantly improves stress discovery accuracy.

III. PROBLEM STATEMENT
Stress is one of the most current health issues in ultra modern society, affecting individuals across all age groups. Nonstop exposure to stress can lead to severe physical and cerebral diseases, including anxiety, depression, and cardiovascular conditions.
Being stress, discovery styles calculate on private approaches such as questionnaires and interviews, which warrant delicacy, thickness, and real-time capability. Also, traditional systems fail to use the full potential of physiological and behavioural data. thus, there is a need to develop an intelligent, automated, and real-time stress discovery system that leverages machine literacy ways to dissect physiological signals and classify stress situations directly.

Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

IV. PROPOSED METHODOLOGY

The proposed system uses a machine literacy- grounded approach for stress discovery and bracket. The methodology consists of the following: .

The proposed system is designed to describe and classify mortal stress situations using machine literacy algorithms grounded on input from physiological and behavioural data. The system workflow begins with data collection from colourful detectors or datasets that capture features such as heart rate, skin temperature, Possible spelling mistake found. Exertion (EDA), facial expressions, or voice tone. After collecting the data, pre-processing ways such as normalization, noise reduction, and point birth are applied to insure data quality and applicability. The gutted and structured data is also fed into a supervised machine learning model — similar as Support Vector Machine(SVM), Random Forest, or Neural Network which is trained to fête patterns associated with different stress situations (low, medium, high). The model is estimated using delicacy, perfection, recall, and F1- score criteria. Once trained, the system can accept real- time or batch inputs to prognosticate an existent's stress position.

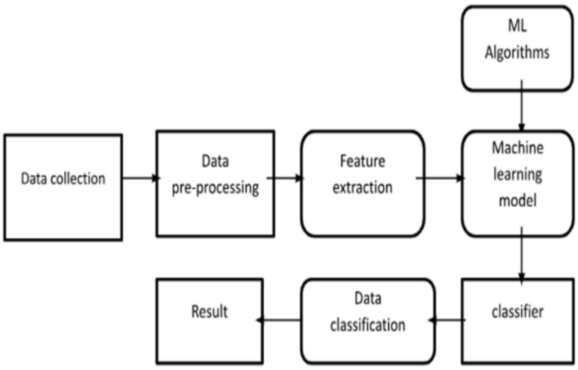

Fig.1. System Architecture

1. Data collection
2. Data pre-processing
3. point of birth
4. Evaluation model
5. Stress bracket
6. System affair

A. Data collection

This step is concerned with opting for the subset of all available data that you will be working If you are writing a formal text, avoid using preposition at the end of sentence.. ML problems start with data rather, lots of data (compliance) for which you formerly knew the target answer. Data for which you formerly knew the target answer is called label data.

B. Data pre-processing

Organize your named data by formatting, drawing and testing from it. Three common data pre-processing ways are Formatting. The data you have named may not be in a format that is suitable for you to work If you are writing a formal text, avoid using preposition at the end of sentence. The data may be in a relational database and you would like it in a flat train, or the data may be in a personal train format and you would like it in a relational database or a textbook train.

Drawing data is the junking or fixing of missing data. There may be data cases that are deficient and don't carry the data you believe you need to address the problem. These cases may need to be removed. also, there may be sensitive information in some of the attributes, and these attributes may need to be anonymized or removed from the data entirely.

Testing: There may be far more named data more available than you need to work If you are writing a formal text, avoid using preposition at the end of sentence.Further data can affect in much longer running times for algorithms and larger computational and memory conditions. You can take a lower representative sample of the named data that may be important faster for exploring and prototyping results before considering the whole dataset.

C. point birth

Coming thing is to do point birth is a trait reduction process. Unlike point selection, which ranks the being attributes according to their prophetic significance, point birth actually transforms the attributes. The converted attributes, or features, are direct combinations of the original attributes. Eventually, our models are trained using the Classifier algorithm. We use classify module on Natural Language Toolkit library on Python. We use the label dataset gathered. The rest of our label data will be used to estimate the models.

D. Evaluation Model

Evaluation is an integral part of the model development process. It helps to find the stylish model that represents our data and how well the chosen model will work in the future. assessing model performance with the data used for training isn't respectable in data wisdom because it can fluently induce overoptimistic and overfit models.

E. Stress Bracket

An essential component of the model development process is evaluation. The trained model is used to classify stress situations into different orders, such as low, medium, and high. The bracket is grounded on the learned patterns from the dataset. The system predicts stress situations when new input data is handed.

F. System Affair

The final affair of the system is displayed through a stoner interface. The prognosticated stress position is shown to the stoner, which helps in covering internal health conditions. The system can be integrated with operations for real- time stress shadowing.

Performance of each bracket model is estimated base on its averaged. The result will be in the imaged form. graphical representation of classified data.

Delicacy is defined as the chance of correct prognostications for the test data. It can be calculated fluently by dividing the number of correct prognostications by the number of total prognostications

V. IMPLEMENTATION:

The Python programming language is used to implement the suggested human stress analysis system because of its ease of use and robust support for machine literacy libraries. The crime is committed in a programming environment that is similar to Jupyter Notebook. Libraries are employed in the system to carry out various functions. Preprocessing and data manipulation are done with NumPy and Pandas. Machine literacy algorithms

like Random Forest and Support Vector Machine (SVM) are applied using Scikit-Learn. Performance analysis and data visualization are done with Matplotlib. Data preprocessing, point birth, model training, and vaticination are among the various modules that make up the system. The dataset is first loaded and gutted by eliminating any incorrect or missing values. To improve model performance, the stripped data is also regularized. Applicable features are identified and used to train the machine literacy models following preprocessing. To estimate the model's performance, the dataset is divided into training and testing sets. In the end, the stress position of fresh input data is predicted using the trained model. The purpose of the experimental setting is to gauge the suggested stress discovery system's performance. Stress-related characteristics, such as physiological signals and behavioral data, make up the dataset used in this study. Usually at a rate of 8020, the dataset is split into two corridor training sets and testing sets. The machine literacy models are trained using the training set, and their performance is estimated using the testing set. The model's efficacy is assessed using performance metrics like delicacy, perfection, recall, and F1-score. These standards aid in determining how well the model categorizes stress levels. Python libraries are used for the perpetration, and every trial is carried out on a system with a standard tackle configuration. The best-performing model for stress vaticination is determined by comparing the outcomes obtained from various methods.

VI. RESULT

Using machine literacy algorithms, the suggested system performs well in identifying and categorizing stressful circumstances. Due to their superior ability to handle intricate and nonlinear data patterns, Random Forest and SVM demonstrate advanced delicacy when compared to Decision Tree The system effectively interprets behavioral and physiological signals to more accurately predict stressful circumstances. Visualization ways similar as graphs and confusion matrices help in understanding model performance

Algorithm	Accuracy	Precision	Recall
SVM	85%	83%	82%
Random Forest	90%	88%	87%
Decision Tree	80%	78%	76%

Fig.2 Result Section

still, the system's delicacy depends on the quality of input data. Noisy or deficient data can affect prognostications. also, real- time perpetration requires effective processing and optimized models. Overall, the results indicate that machine literacy provides a dependable and scalable result for stress discovery. outcomes.

VII. CONCLUSION

Human stress analysis using machine literacy presents a promising result for accurate, real- time discovery and bracket of stress situations. By using physiological and behavioural data, the system can identify stress patterns and give timely perceptivity that support internal health and well- being Stress Bracket's trustworthiness is increased by the use of supervised literacy algorithms, and its integration with stoner-friendly interfaces makes it usable on a daily basis. The advantages of early stress detection and visionary operation outweigh the drawbacks, notwithstanding issues with data quality, sequestration, and processing loads. This design lays the groundwork for more intelligent, technologically advanced methods of internal health monitoring.

VIII. FUTURE WORK

There are a number of approaches to improve the performance and practicality of the suggested system. To improve stress detection accuracy, future research will use deep learning methods like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). Additionally, wearable technology like fitness trackers and smartwatches can be integrated into the system to enable real-time monitoring. Larger and more varied datasets can also be used to enhance the model's generalization and dependability. Prediction accuracy can be further improved by using multimodal data, such as physiological signs, voice analysis, and facial expressions. To guarantee the secure management of sensitive user data, privacy and security measures might also be reinforced.

Additionally, creating a web-based or mobile application can improve the system's usability. The application of Explainable Artificial Intelligence (XAI) approaches can enhance transparency and assist consumers in comprehending the process of stress prediction.

Acknowledgment

The administrators and teachers of the Department of Computer Science at Jeppiaar University are much appreciated by the writers for their unwavering support and direction during the project's development. Additionally, we would like to express our gratitude to our project guide for their insightful advice and support, which enabled us to successfully finish this work.
Lastly, we would want to express our gratitude to our friends and family for their encouragement and support throughout this effort.

REFERENCES:

1. Beheshti-Kashi, S., Karimi, H.R., Thoben, K.D., Lutjen, M., and Teucke, M. "A check on retail deals soothsaying and vaticination in fashion requests," Systems Science & Control Engineering 3(1), 154, 161 (2015).
2. Smith, Oliver, and Thomas Raymen. "Violent Black Friday shopping in the British context" Consumer Culture Journal, 17.3 (2017), 677-694.
3. 3)Goutam Majumder. "ANALYSIS AND VATATION OF CONSUMER GEST ON BLACK FRIDAY DEALS" Gujarat Research Society Journal 21.10 (2019) 235-242.
4. "Empirical assessment of machine literacy grounded software disfigurement vaticination ways" by Challagulla, Venkata Udaya B., et al. Artificial Intelligence Tools International Journal, 17.02 (2008), 389-400.
5. Chu, C.W., Zhang, G.P. "A relative study of direct and nonlinear models for aggregate retail deals soothsaying," International Journal of Product Economics 86(3), 217{231(2003)}.
6. Makridakis, S., Wheelwright, S.C., and Hyndman, R.J. "Soothsaying styles and operations," John Wiley & Sons (2008).
7. Cassio P. de Campos, Robert Peharz, Correia, and Alvaro. "Joints in Random Timbers" Neural Information Processing Systems Advances 33 (2020).
8. Kvalheim, Olav Martin, et al. "Determining the optimal number of factors in partial least places retrogression from distributions of the root mean squared error obtained by Monte Carlo e testing" Chemometrics Journal 32.4 (2018) e2993.
9. "Extreme grade boosting as a system for quantitative structure – exertion connections" by Robert P. Sheridan et al. 56.12 (2016) 2353-2360 Journal of Chemical Information and Modeling.
10. Wei Khor, Ngiam, and Kee Yuan. The Lancet Oncology 20.5 (2019) e262-e273.
11. Domingos, P.M. "Big data and machine literacy algorithms for health-care delivery." There are numerous benefits to understanding machine literacy. Communacm 55(10), 78–87 (2012).
12. Langley, P., Simon, H.A. Machine literacy and rule induction applications. ACM Dispatches 38(11), 54(64) (1995).
13. On September 20, 2020, the website https://machinelearningmastery.com/gentle-introductionxgboost-applied-machine-learning was accessed.